

Review

Methods for assessment of keel bone damage in poultry

T. Casey-Trott,^{*} J. L. T. Heerkens,[†] M. Petrik,^{*} P. Regmi,[‡] L. Schrader,[#] M. J. Toscano,^{§,1}
and T. Widowski^{*}

^{*}Animal and Poultry Science, University of Guelph, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada;

[†]Institute for Agricultural and Fisheries Research, Animal Sciences Unit – Farm Animal Welfare & Behaviour Research Group, Scheldeweg 68, 9090 Melle, Belgium; [‡]Department of Animal Science, Anthony Hall, 474 S.

Shaw Lane, East Lansing, MI 48824 USA; [#]Institute of Animal Welfare and Animal Husbandry,

Friedrich-Loeffler-Institut, Doernbergstrasse 25/27, 29223 Celle, Germany; and [§]Research Center for Proper Housing: Poultry and Rabbits (ZTHZ) Division of Animal Welfare, VPH Institute, University of Bern,

Burgerweg 22, 3052 Zollikofen, Switzerland

ABSTRACT Keel bone damage (KBD) is a critical issue facing the laying hen industry today as a result of the likely pain leading to compromised welfare and the potential for reduced productivity. Recent reports suggest that damage, while highly variable and likely dependent on a host of factors, extends to all systems (including battery cages, furnished cages, and non-cage systems), genetic lines, and management styles. Despite the extent of the problem, the research community remains uncertain as to the causes and influencing factors of KBD. Although progress has been made inves-

tigating these factors, the overall effort is hindered by several issues related to the assessment of KBD, including quality and variation in the methods used between research groups. These issues prevent effective comparison of studies, as well as difficulties in identifying the presence of damage leading to poor accuracy and reliability. The current manuscript seeks to resolve these issues by offering precise definitions for types of KBD, reviewing methods for assessment, and providing recommendations that can improve the accuracy and reliability of those assessments.

Key words: Hen, bone, welfare, damage, keel

2015 Poultry Science 94:2339–2350

<http://dx.doi.org/10.3382/ps/pev223>

INTRODUCTION

Keel bone damage (KBD) is a major problem for commercial laying hens, with the United Kingdom's Farm Animal Welfare Council (FAWC) advising that it is one of the most important issues facing the industry (FAWC, 2010, 2013). The primary concern stems from the pain birds are likely to experience following KBD which has critical implications for welfare (Nasr et al., 2012b). Additionally, the limited information that exists suggests that KBD negatively affects productivity (Nasr et al., 2012a), posing financial concerns for producers. In response to this problem, multiple research efforts are focused on assessing the causes of KBD and associated risk factors, the consequences for the birds, economic impacts, and means to reduce their occurrence and severity. Current approaches for studying KBD were recently reviewed at the 2014 International Keel Bone Damage Workshop and recommendations for research to determine the causes, solutions, and impli-

cations for welfare and productivity in hens with KBD were agreed on and have recently been published elsewhere (Harlander-Matauschek et al., 2015). The present paper focuses exclusively on methods of assessing KBD and their appropriateness given different environments, research objectives, and budget constraints. We initially provide definitions for the types of KBD typically seen in commercial settings and then review a variety of methodologies to assess damage. Of the discussed methods, palpation is by far the most common technique used to assess KBD, but its accuracy, as well as the ways in which results are reported, vary considerably across studies, limiting its usefulness as a research tool for comparison in different contexts. Both the quality and reliability of KBD assessment by palpation could benefit from individual assessors improving their abilities at identifying damage. To that end, we provide suggestions for improving detection with palpation, review the literature on training methods that have proven successful in related fields, and identify potential sources of error in assessment. We also review specific criteria for reliability and consistency that assessors should achieve in order to consider themselves competent. Finally, we recommend a format for recording and reporting KBD to improve comparison across studies.

© 2015 Poultry Science Association Inc.

Received January 28, 2015.

Accepted June 15, 2015.

¹Corresponding author: Michael.toscano@vetsuisse.unibe.ch

DEFINING KEEL BONE DAMAGE: FRACTURES AND DEVIATIONS

The Keel Bone

The keel is a pronounced bone that extends from the sternum and runs axially over the midline. It is situated ventral to the heart where it anchors the muscles used for wing motion, the *pectoralis major* and *pectoralis minor*. The length of a keel bone from the *Carina apex* to the caudal tip along the ventral surface (Figure 1) is approximately 9 to 12 cm, though this varies with genetic line, age, and other factors (Toscano, personal observations). The height of the keel measured from the *Carina apex* to the dorsal surface is approximately 30 to 33 mm (Casey-Trott, Heerkens, Toscano, personal observations). Damage can broadly be divided into 2 categories: fractures and deviations.

Fractures

Fractures are characterized by sharp bends, shearing, and/or fragmented sections of the keel bone. Fractures may extend from the ventral to the dorsal surface in the sagittal plane, though they can also be cranial to

caudal, or a combination of both. Photographs of fracture damage with increasing severity in adult laying hens can be viewed in Wilkins et al. (2011) and radiographic images of the progression of healing are given in Richards et al. (2011). The possible causation of keel fractures is unknown, though collisions with internal housing elements have been suggested as a potential cause. The role of collisions as well as other factors as a cause of fractures (e.g., strong muscular contractions) are discussed elsewhere (Harlander-Matauschek et al., 2015; Sandilands et al., 2009).

A keel fracture in a live bird is most typically identified (through palpation) by the presence of callus material on the ventral and lateral surfaces as a product of the regenerative healing process in the period after the fracture has occurred. The healing process consists of several phases each involving the formation of a different type of tissue (Einhorn, 2005). The initial phase of the regeneration process starts with an inflammatory response that results in the formation of a hematoma and granulation tissue. Next follows the formation of a soft callus on the fracture site consisting of cartilaginous or chondroid tissue. In the third phase, ossification of the soft callus takes place to form a hard or bony callus (Einhorn, 2005) consisting primarily of woven bone

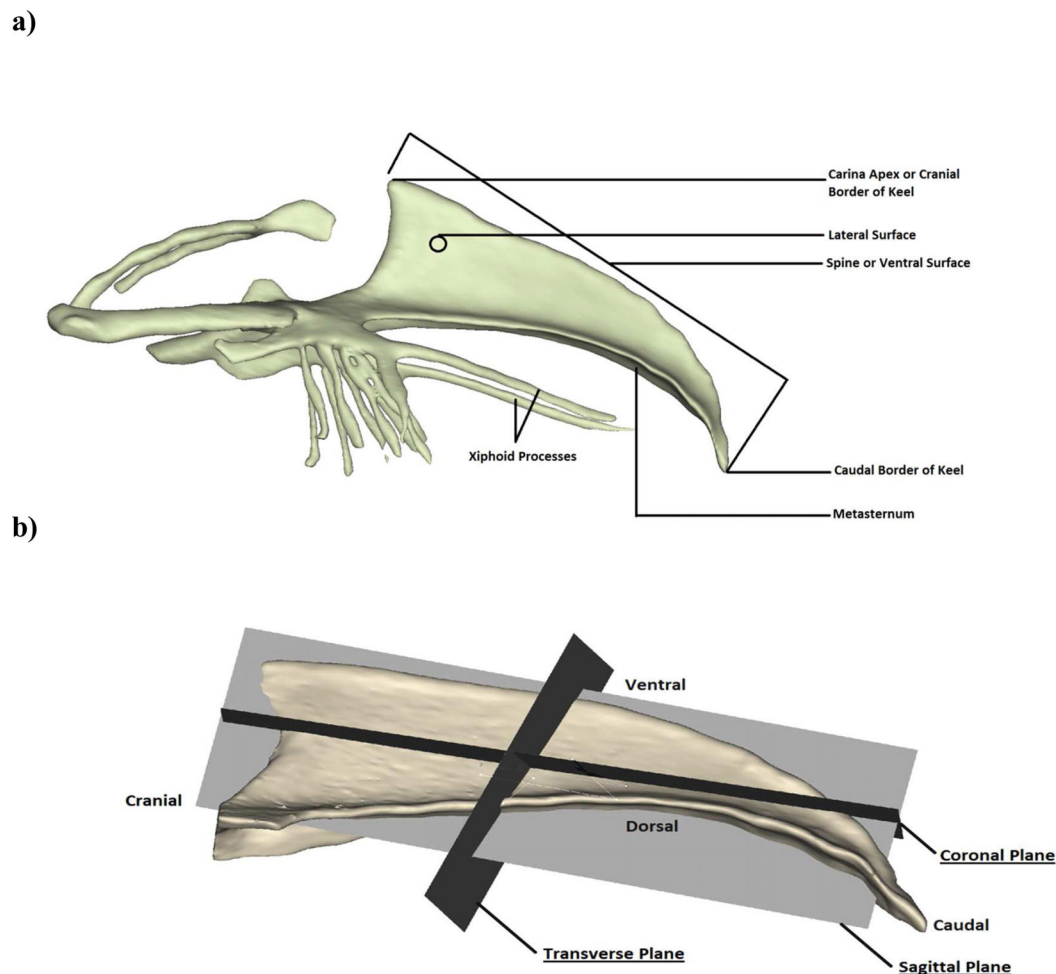


Figure 1. Keel anatomical descriptions (a) and orientations (b).

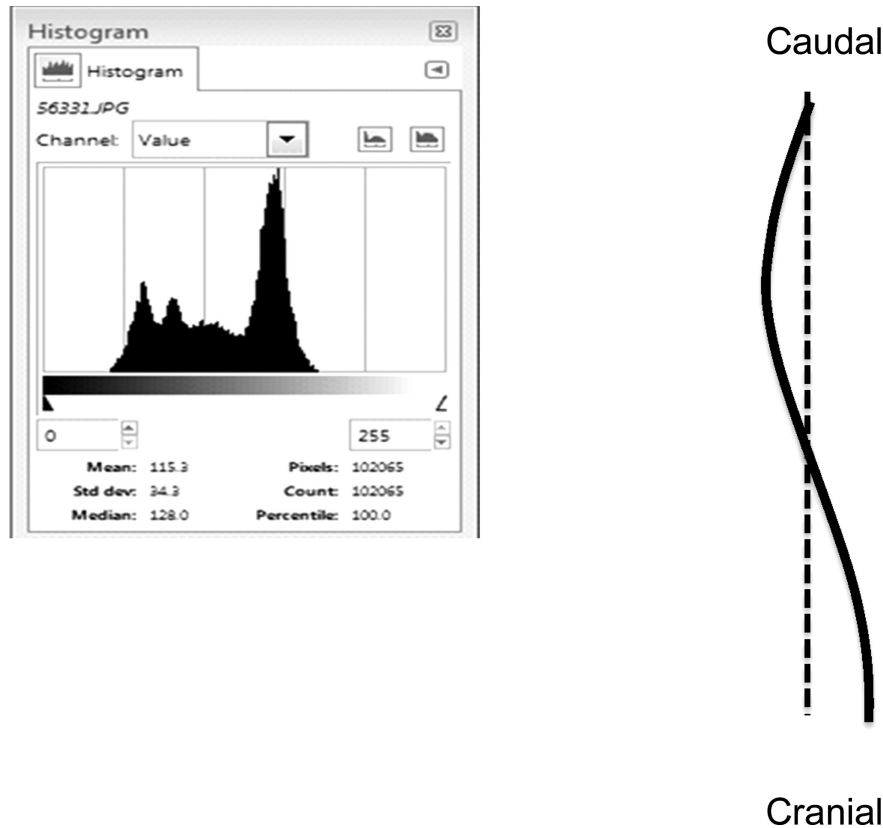


Figure 2. Method of quantifying deviation from a straight line. Viewing the keel with the ventral surface facing towards the observers (where the ventral surface is represented by the solid line), software can be used to calculate the area or number or pixel numbers from a line drawn between distal and caudal points on the keel (represented by the dashed line). Output from software calculating the pixels associated with the variation is shown.

tissue. Eventually, the woven bone is gradually replaced by lamellar bone tissue during the remodeling phase.

The presence of a callus can be fairly obvious by palpation, though this will vary with the size and location of the fracture, and in minor cases it will be relatively small (i.e., $<1 \text{ mm}^3$) and thus more difficult to detect. For the callus characteristics to appear, the fracture must be old enough for it to have formed, and hence they are often referred to as ‘old breaks’ (Wilkins et al., 2004). As an additional characteristic, a bulbous area around the fracture site resulting from inflammation can indicate the presence of recent or ‘new’ fractures. Surprisingly, evidence of new fractures is rarely found (Toscano, personal observations) suggesting this inflammation may occur only for a short time. In conclusion, we propose that keel fractures be defined as keels manifesting the presence of a periosteal scar or callus which would indicate the ongoing healing process following a fracture. Presence of the inflammatory response should also be included, though ideally noted separately.

Deviations (from normal)

While fractures are characterized by a defect in the architecture of the bone and an interruption in the periosteum, a second type of KBD commonly seen lacks

these characteristics and requires its own definition. Terms to describe this alternate damage have included: bending, S-shaped, twisted, or curved keels (Fleming et al., 2004; Lay et al., 2011; Habig and Distl, 2013), though these can be easily misinterpreted without a visual image and thus we advocate a more descriptive definition. The possible causation of this class of damage and its implications for animal welfare are discussed elsewhere (Harlander-Matauschek et al., 2015; Sandilands et al., 2010; Pickel et al., 2011). We propose that a keel bone deviation be defined as a bone with an abnormally shaped structure that has not resulted from a fracture but contains section(s) that vary from a theoretically perfect 2-dimensional straight plane in either the transverse or sagittal planes. Additionally, indentations along the ventral surface can also be classified as a deviation.

METHODS FOR ASSESSING KEEL BONE DAMAGE (OTHER THAN PALPATION)

Recent efforts have sought to expand the methods which can be used to assess KBD beyond palpation and visualization with the naked eye. While we believe that palpation will continue to be the chief means of assessing KBD due to its low cost, ease of adoption, and

the method being well-validated, alternative methods should be considered as they can offer additional insight that is not possible by palpation alone. In this section, we provide a short summary of non-palpation methods to assess damage as a resource for researchers interested in identifying a technique that meets their particular needs. The reviewed techniques are not exhaustive, but include methodologies that range from relatively 'low tech' that can be easily performed on-farm at minimal cost, to methods that require access to expensive equipment and specialized facilities (e.g., magnetic resonance imaging). All of the discussed methods can be performed on live birds and thus provide several benefits including capacity for longitudinal observations, reduced loss of egg production, and an overall decrease in the loss of life. Euthanizing the bird and a visual inspection with the 'naked eye' would allow for a relatively objective assessment of damage, though has obvious detriment. Additionally, even the inspection of an excised keel may fail to detect internal damage that some of the methods described in this section would identify. Each of these concerns would need to be evaluated to determine the most appropriate method for the particular circumstances and goals of a given experiment. The techniques have different strengths, and with each we have tried to present the relative benefits and detriments.

X-radiation (X-ray)

A form of electromagnetic radiation, X-rays can be directed into and through the body towards a detector located behind the image of interest. Due to variation in the absorbance capacity of various tissues (e.g., bone has a relatively high concentration of calcium which absorbs radiation effectively), an image or 'shadow' is created on the detector which can indicate patterns of structures within the body. The technique is limited in that only 2-dimensional images can be produced. Additionally, repeated X-ray exposure can be a potential health hazard to the bird, and the human operator, but reasonable precautions can minimize this risk. Radiographic equipment is widely available, particularly in universities and veterinary clinics, and has been successfully used to assess keel damage in anesthetized birds (Richards et al., 2011). More recently, work by Sirovnik et al. (*in prep*) developed a procedure using portable X-ray equipment that allowed for images to be made directly in the production facility without anesthesia by hanging the birds upside down and exploiting their natural inclination to remain still in this position. For the study, only 6 birds at a single age were used, with varying levels of keel fracture, thus additional work is needed to validate the method. One of the advantages of using radiographs is that it allows identification of recent fractures (e.g., before an inflammatory response) and fractures involving the dorsal aspect of

the keel, which are impossible to identify by palpation alone (Richards et al., 2011).

Ultrasonography

Using high frequency sound (ultrasound) waves, ultrasonography captures the reflection of sound from structures within the body yielding images of various tissues. A major benefit of ultrasound is that neither the human operator or hen are exposed to ionizing radiation and thus the technique is much safer than the X-ray approach. The technique has been used to assess keel fractures in adult hens (Sandilands et al., 2010), although it has not been extensively validated. In this method, the associated probe is run along the ventral ridge of the keel bone and the recorded image is observed for fractures. Depending on the breast muscle mass, the probe can also be run along the lateral surface of the keel to obtain additional information on deviations. The biggest challenge in using ultrasonography is determining the appropriate size and shape of the probe to best detect fractures, although the associated equipment is likely to be far less expensive than that for the X-ray approach, and safer for the operator. The feather cover around the breast, especially in pullets and young hens, as well as breast muscling, could also impact the efficacy of this technology. Lastly, it is likely that it will be more difficult to identify damage in the generated images than the other imaging techniques discussed (i.e., radiography and tomography).

Peripheral quantitative computed tomography (pQCT)

Tomography uses a narrow beam of low level radiation and clinical computed tomography machines can conduct a full scan in few minutes. Depending on the gantry size of the scanner, multiple hens can be scanned at the same time. As the scan time is very short, hens can be restrained manually without use of any anesthetics. Resulting 2-dimensional images can be post-processed with imaging software to create 3-dimensional models of the keel bone which can be used for further calculations, e.g., average volumetric density at any particular cross-section of the keel bone (Regmi et al., 2013). The final images are expressed in Hounsfield units and common imaging software like MIMICS (Materialise, Plymouth, MI) enable the operator to select an appropriate Hounsfield unit threshold to separate surrounding soft tissues from the keel bone. The resulting 3-dimensional images are of a highly accurate geometry. A major advantage of pQCT over the other described methods is that the 3-dimensional image can be rotated in all planes (360 degrees) to identify minor breaks and fracture calluses. The technique is also of particular benefit for identifying fractures on the dorsal surface and the tip of the keel which are very difficult to identify with other techniques. As a

drawback, pQCT machines used for clinical purposes generally have a large voxel size which makes it impossible to differentiate the various types of tissue within the keel. Also, machines are not portable and thus cannot be used for on-farm situations.

Straight line reference

To quantify deviations in the sagittal plane, Stratmann et al. (*personal communication*) quantified the amount of deviation in excised keels by comparing the actual path of the keel's ventral surface with a referential straight line in terms of the area within the boundary created by these 2 lines. The study was conducted as a preliminary investigation and involved only 10 end of lay, Lohmann Selected Leghorn hens with a range of deviated keels, thus the procedure will require further study to ensure validity. Although the method does require a computer and relevant free software (e.g., ImageJ available from National Institute of Health), analysis could theoretically be performed with photographs and thus allow for both on- and off-farm assessments.

Future Technologies

New, non-invasive, technologies need to be developed that identify chickens with keel fractures without handling them. The possibility exists that handling itself could contribute to keel bone fractures, especially in aging laying hens as they have to be captured, assessed, and then released during the evaluation process (Knowles and Wilkins, 1998). Potential technologies include automated images of keels that assess changes in electrical or thermal properties of the bone that correlate with a fracture. Thermography of the injury site has been shown differences between hens with and without fractures, and unexpectedly suggested that the keel area was cooler when fractures were present (Nasr et al., 2012a).

ASSESSING KEEL BONE DAMAGE BY PALPATION AND SUGGESTIONS FOR IMPROVEMENT

Currently, the most common method of evaluating keel damage in conscious birds is by palpation due to its simplicity and relatively low cost. Despite its wide use, palpation for keel fractures is likely to suffer from poor accuracy and repeatability, particularly when assessors have not undergone appropriate training. The following section reviews several recommendations to improve the quality of assessment and provide guidance to those performing palpations.

Diagnostics to Assess Validity and Reliability of Palpation

Palpation is a manual technique that is used for the localization and confirmation of anatomical landmarks such as bone, muscles, and organs, or to derive information about function, such as pulse rate (Ullrich and Kuhlen, 2012). The procedure is widely used within the fields of human and veterinary medicine as a diagnostic tool for identifying injury, or as a method to determine the necessity for more comprehensive assessments. Although palpation is routinely used by practitioners, the validity and reliability of associated techniques have historically been a point of contention (Seffinger et al., 2004). To evaluate the effectiveness of a methodology in identifying whether individuals are affected by the condition, quantification of sensitivity, specificity, and the positive and negative predictive values should be assessed. A detailed explanation of these terms and how they are calculated can be found in McKenna and Dohoo (2006). Reports of low sensitivity using palpation to detect certain conditions (Wiest et al., 1998), as well as low intra- and inter-observer reliability (Seffinger et al., 2004; Degenhardt et al., 2005; Wood et al., 2006; Myers et al., 2011), indicate that a detailed protocol and comprehensive training both play a vital role in determining whether a palpation method is capable of producing accurate, precise, and consistent results.

In addition to quantifying the attributes of diagnostic tests to gauge the quality of collected data (Dohoo et al., 2009), it is equally important to know the true status of a condition via a gold standard (McKenna and Dohoo, 2006). Subjective visual determination of fractures, following post-mortem dissections, has generally been used as the "gold standard" to determine the true prevalence of fractures and the specificity, sensitivity, and accuracy of palpation for assessing keel damage. However, few studies have assessed the correlation between scores from gross visual assessment of dissected keels, and more accurate measures such as radiographic or microscopic examinations. Some keels identified as non-fractured during post-mortem visual examination have been shown to have histological evidence of fracture (Scholz et al., 2008), suggesting even the gold standard currently in use may be questionable. Richards et al. (2011) found that radiography was more sensitive than palpation for detecting minor fractures in live birds, with new and/or minor fractures characterized by a radiolucent fracture line. It is not however known whether these minor fractures could also be detected by the naked eye. It may be best at this time to consider values for keel fractures determined from post-mortem visual assessment as "apparent" prevalence rather than "true" prevalence.

Additionally, sensitivity and specificity of a test can vary across populations (Dohoo et al., 2009). To date, values for sensitivity and specificity of palpation techniques based on comparisons with visual assessment of dissected bones have only been reported for keel and

furculum fractures in end of lay hens (strains not reported; Wilkins et al., 2004) and for keel fractures in brown hens near end of lay (Petrik et al., 2015). When evaluating furcular and keel fractures, Wilkins et al. (2004) reported an accuracy of > 70%, and sensitivity between 0.86 and 0.87. Specificity, on the other hand, ranged from 0.38 to 0.76 indicating a high degree of false negatives which the authors attributed to a varied level of experience among the examiners. Similar ranges were reported by Petrik et al. (2013) where only keel bones were scored: accuracy 87 to 97%, sensitivity 0.72 to 0.87, and specificity 0.52 to 0.67. The apparent prevalence of fractures, confirmed by dissection, was similar in both studies at 65% in Wilkins et al. (2004) and 62% in Petrik et al. (2013).

In addition to the test attributes discussed above, the intra- and inter-rater reliabilities of a test (consistency or level of agreement in scoring within and between observers, respectively) are critical for determining the quality of data generated from its application (Viera and Garrett, 2005). Cohen's Kappa statistic is generally used for testing levels of agreement above that expected by chance and its value can be used to interpret how well raters agree (Landis and Koch, 1977). Additionally, having observers blind to experimental treatment may be important for reducing observer bias which can lead to error in scoring (Hróbjartsson et al., 2013). In relation to the current topic, a recent report on the effect of soft perches on KBD determined intra-observer reliability by having 2 observers involved in the study score a sample of live birds twice, with a 2-hour interval between scorings (Stratmann et al., 2015b). The investigators subsequently evaluated inter-observer agreement on 43 live birds following completion of the experiment. They reported Kappa values of 0.7 (substantial) and 0.54 (moderate) for intra- and inter-observers reliabilities, respectively. These authors also reported conformance between palpation and post-mortem assessment ($\kappa = 0.53$) with an accuracy of 68%. Self-evaluation such as these should be considered as a requirement when reporting KBD.

Training for Palpation

The primary approach to improving palpation is to better develop the skills and consistency of the examiner. Improvements can be achieved through a variety of means, including technologically advanced solutions used during the training phase to evaluate and improve the technique of the examiner. Some examples include using 3-dimensional animations (van Sint Jan, 2007; Norman and Dall'Alba, 2013) or computed tomography scans (McCormick et al., 2003) to improve the visualization of anatomical structures. Medical simulations with virtual palpations (Bossaert et al., 2009; Ullrich and Kuhlen, 2012), sensor-enabled mannequin technology (Pugh, 2013), or force-feedback technology (Baillie

et al., 2003) have also been used to train or evaluate palpation skills.

When technologically advanced methods are not possible, targeting how palpation techniques are best learned and understood can lead to better training methods. Palpation training is a critical part of the curriculum for veterinarians, doctors, physiotherapists, and osteopaths and yet it is often referred to by students as one of the most difficult clinical skills to learn (Esteves and Spence, 2014). To become an expert at a given task, thousands of hours of practice of the technique are required (Ericsson et al., 2007). In training of veterinary students, for example, the Australian Cattle Veterinarians recommend that 2,000 cows be palpated for pregnancy before a student is considered to be eligible to take their competency exam for this procedure (Norman and Dall'Alba, 2013). In terms of palpation expertise, although gender and level of medical education did not affect accuracy, higher overall experience level significantly improved the validity and reliability of palpation techniques used by human medical practitioners (Pugh, 2013).

Esteves and Spence (2014) proposed several steps for palpation training of osteopathic students. According to their recommendations, meaningful practice is required to consistently locate anatomical landmarks, and skills can be improved by having a thorough understanding of the anatomical features, physiologic function, and biomechanical role of the region of interest. Additionally, exercises that encourage visuo-spatial thinking train the brain to more effectively use tactile skills to produce a mental 3-dimensional image, and enhance the examiner's ability to visualize the structure. Once a thorough understanding of the form and function of the structure has been grasped, palpation in damaged, altered, or imperfect keels should then be included in the mix of practice palpations. Discussions regarding the causality and probability of the injury, as well as providing feedback and interpretations of the student's skills were recommended as a regular part of the training process. Regardless of level of expertise, Esteves and Spence (2014) considered routine reflection on one's own validity and reliability in practice as essential to developing and maintaining competent palpation skills.

Learning the skill of palpation provides the foundation for producing accurate palpation diagnoses; however, maintaining consistent results is a constant challenge. Degenhardt et al. (2005) reported improvements to intra- and inter-observer reliability when consensus training was used. Consensus training allows all examiners to observe each other's individual palpation and scoring techniques. The group then discusses which techniques most effectively located the anatomical features and produced an accurate diagnosis, followed by the creation of a detailed description of the agreed on palpation protocol that will be used by the group. A second study also recommended routine 'recalibration', which is a process in which the individual is

reminded of the original palpation protocol and subsequently assesses their current techniques and makes the necessary adjustments to return to the original protocol (Degenhardt et al., 2010).

To date, Petrik et al. (2013) is the only study that reported the accuracy and inter-observer reliability of scoring palpations for keel fractures in laying hens over the course of training. In that study, 8 assessors scored a total of 100 laying hens and achieved an overall mean accuracy of 91.8%, Kappa value of 0.437, and a true prevalence of 67% determined by dissection. The training process involved a practice session on 10 laying hens to familiarize the examiners with the technique, followed by individual scoring and an assessment of the accuracy and inter-observer reliability on those 10 birds through consensus training during post-mortem dissection and examination, an approach similar to that suggested by Degenhardt et al. (2005). The assessors then independently scored an additional 90 birds. Mean accuracy of individual assessors increased significantly from birds 1 through 10 (73.3%) to birds 11 through 100 (94.1%). The Kappa value from the first 50 birds was 0.414, whereas the Kappa value for the last 50 birds was 0.474 (Petrik et al., 2013) indicating that consensus training and practice did improve the inter-observer reliability of the group. The results from this study provide useful reference values and an example of a successful training method. More studies of this kind are warranted to establish a repeatable range of Kappa values and test ways to further improve the inter-observer reliability of the examiners.

Variations in the Criteria for Assessing KBD

The palpation methods currently in use originate primarily from 2, key seminal works (Wilkins et al., 2004; Scholz et al., 2008) which have since been adopted and/or combined by a variety of research groups in their assessment of KBD. While broadly similar in their classification of damage, a key difference exists and requires resolution. In contrast to Scholz et al. (2008), Wilkins et al. (2004) did not incorporate the existence of deviations (bends in the transverse and sagittal planes). Thus, this initial work by Wilkins et al. (2004) and subsequent assessments by the Bristol group (Richards et al., 2011, 2012; Wilkins et al., 2011; Tarlton et al., 2013; Toscano et al., 2013; Nasr et al., 2015) and others (Rodenburg et al., 2008; Nicholson and O'Connell, 2010; Donaldson et al., 2012) based their grading of fractures independent of deviations. Keel bones with deviations were observed in developing the Wilkins method; however, they were excluded from the scoring system as this alteration in shape was believed to result from a normal remodeling process in response to continual, low-grade pressures. For instance, perching is a likely source of such pressure and has been shown to be associated with deviations (Tauson and

Abrahamsson, 1994; Vits et al., 2005; Barnett et al., 2009; Pickel et al., 2011; Regmi and Karcher, 2013).

In contrast to the Wilkins/Bristol-developed method focusing entirely on fractures, Scholz et al. (2008) incorporated deviations into their assessment as well as fractures. Using 162 keels from euthanized hens, Scholz et al. (2008) found that 51% of keel bones which showed a slight deformity macroscopically (i.e., no visible fracture seen by the naked eye but some type of deviation was present) had histological evidence of a fracture in the form of newly woven bone. Moreover, the newly woven bone was comparable to the bone growth present on more severe macroscopic deformities, leading them to suggest that this subgroup of 'slight deformities' had actually suffered a fracture as well, a finding also reported by others (Fleming et al., 2004). The results by Scholz et al. (2008) strongly indicate that even slight deformations which are not recorded in the scheme by Wilkins et al. (2004) may actually possess keel bone fractures in approximately half of those identified as lacking damage. Moreover, a scoring scheme based on palpation of living hens might underestimate even moderate and severe deformations in comparison to a more thorough visual assessment of carcasses. These concerns have led other research groups (Kappeli et al., 2011; Gebhardt-Henrich and Fröhlich, 2013; Heerkens et al., 2013; Habig and Distl, 2013; Stratmann et al., 2015a, b) to incorporate keel bone deviations into their scoring systems. The subsequent variety in scoring systems and methodologies has led to challenges in comparison and interpretation of results from different groups.

A Binary System to Improve Palpation Assessment

In identifying the most appropriate means to assess KBD, different methods exist, with each offering particular benefits and shortcomings in usage. Regarding palpation, given the complications in making comparisons across studies, the authors propose a simplified scoring system which would be beneficial to the collective effort of identifying the causes of KBD and reducing their occurrence and severity. In addition to the direct benefit of greater ability to compare across studies, a simplified system allows for several indirect benefits including:

- Improved comparisons across research groups leading to greater consistency and less replication, a key component of the 3R's objective for ethical research (<https://www.nc3rs.org.uk/the-3rs>).
- Easier interpretation of results by government and non-governmental agencies in efforts to formulate and develop recommendations and policy.
- Possibilities for meta-analysis across types of systems, feeding and management practices, and genetic lines.

Our proposed system – the Simplified Keel Assessment Protocol (**SKAP**) – has several features which we believe will ease reporting, interpretation, and comparison when assessing KBD in live hens. Firstly, SKAP addresses both fractures and deviations, but as 2 separate, mutually exclusive, binary categories where a total of 4 resultant possibilities exist. It is unknown if keel deviations without fractures are associated with pain or affect the animal's welfare, though it is possible that deviations weaken the structure of the keel, rendering it more susceptible to fracture. In support of this notion, White Leghorn hens at 71 wk of age with severe keel deformities had poorer bone mineral density compared to hens with normal keels (Hester and Enneking, 2014). A weakened structure is particularly concerning given that half of keels with slight deviations in one study contained histological evidence of fracture (Scholz et al., 2008). Thus we believe deviations should be included in assessments given the possibility that they may have a direct effect on welfare as well as indirectly contributing to the occurrence of fractures.

Secondly, the SKAP system does not attach a measure of severity to either deviations or fractures, but rather their presence or absence. We believe this to be a benefit for investigations as it eliminates the difficulties and ambiguities in determining whether some fractures are less or more severe than others, and in comparing these categories across efforts in different labs. The criteria for assigning severity grades to KBD varies across research labs and include: the size of the fracture, whether deviations are present, and categories such as 'slight' or 'major' which are not explained further and thus offer poor objectivity. Percentages within these scoring systems are often provided (Kappeli et al., 2011; Wilkins et al., 2011; Toscano et al., 2012; Tarlton et al., 2013) and can suggest patterns over time, though will have little meaning outside the research group of the authors. While we do believe that severity measures would be an important tool in determining when damage becomes severe enough to warrant intervention, the authors concluded that information to establish meaningful thresholds for severity grades is currently lacking. In order to make definitive distinctions in regards to lesser or greater severities of KBD, we encourage researchers to develop and validate such methods and incorporate them into the SKAP methodology. For instance, results of severity can and should be reported alongside the binary classifications of fracture and deviation, which practice would not take away from the simplified format and the discussed benefits. In this manner, the presence or absence of KBD can be reported and interpreted easily with minimum ambiguity. Ideally, reports containing efforts at grading severity can be applied retrospectively to data previously collected using the SKAP format. Given this possibility and the numerous academic research groups, institutions, and certification schemes that are conducting assessments of KBD, the authors envision potential for a web-based platform where scientists could upload their results

Table 1. Prevalence of keel bone damage scored using the Simplified Keel Assessment Protocol (SKAP) with data from both palpation of live birds and visual assessment during dissection (n = 202). Absolute numbers as well as percentages of the total in parentheses are provided.

| Assessing Keel Bone Damage by palpation | | | | |
|--|-------|-----------|-----------|------------|
| Deviation | | | | |
| | | Yes | No | Total |
| Fracture | Yes | 71 (35%) | 35 (17%) | 106 (52%) |
| | No | 22 (11%) | 74 (37%) | 96 (48%) |
| | Total | 93 (46%) | 109 (54%) | 202 (100%) |
| Assessing Keel Bone Damage by dissection | | | | |
| Deviation | | | | |
| | | Yes | No | Total |
| Fracture | Yes | 70 (35%) | 43 (21%) | 113 (56%) |
| | No | 37 (18%) | 52 (26%) | 89 (44%) |
| | Total | 107 (53%) | 95 (47%) | 202 (100%) |

using a SKAP scoring sheet into a database which could be publicly accessed and used for analysis by others. To facilitate and encourage the uptake and use of the system, we have developed a scoring sheet which is publicly available via the provided internet link¹ or by contacting the corresponding author.

Understanding Errors in Palpation Scoring Using SKAP

The levels of accuracy, sensitivity, and positive predicted values (true positives/no. of positive calls × 100) discussed earlier indicate that palpation is an effective method to detect the presence of fractures more than 70% of the time; however, approximately 30% of fractures (as determined by a visual inspection at dissection) can go undetected. While some of those fractures may be impossible to detect by palpation alone, it is possible that a portion of false negatives could be eliminated by more intensive training, practice, and an understanding of where errors commonly occur.

A total of 202 live end of lay (68 to 81 wk old) Lohmann LSL Lite hens were palpated and scored by a single investigator. Birds were selected for scoring by live palpation from furnished cages, conventional cages, and an aviary system and subsequently euthanized for post-mortem assessment to determine the true nature of the fractures and deviations. The observer was blind to the palpation keel score during the post-mortem assessment of the true keel status. Complete results of the outcome are provided (Table 1). Criteria for KBD (fractures and deviations) were those proposed in this manuscript using SKAP, though we attempted to include a measure of severity for deviations including: no deviations, deviations <0.5 cm from 180° in the sagittal plane (minor), or greater (severe).

The overall accuracy for detecting fractures by palpation was 84%, with sensitivity 81%, and specificity

¹<https://drive.google.com/file/d/0B0SUUWV0oDdQakxNUzY5MlhKMmc/view?usp=sharing>

Table 2. Detailed description of the location of fractures that were correctly assessed by palpation (True +), as well as location of fractures that led to errors in palpation (False - and False +) as compared to the true nature of the keel at dissection. A total of 73 keels were correctly scored as having no fracture (True -). Accuracy of fracture detection by palpation was 84%. A total of 202 keels were assessed. Absolute numbers as well as percentages of the total in parentheses are provided.

| Fracture Location | True + | False - | | False + | |
|-------------------|----------|-----------|------------|-----------|------------|
| | | Palpation | Dissection | Palpation | Dissection |
| No Fracture | N/A | 22 | 0 | 0 | 11 |
| Tip | 58 (60%) | 0 | 18 (82%) | 7 (63%) | 0 |
| Middle | 8 (8%) | 0 | 0 | 2 (18%) | 0 |
| Tip & Middle | 30 (31%) | 0 | 4 (18%) | 2 (18%) | 0 |
| Total | 96 (48%) | 22 (11%) | | 11 (5%) | |

Table 3. Detailed description of the severity of deviations that were correctly assessed by palpation (True +), as well as location of deviations that led to errors in palpation (False - and False +) as compared to the true nature of the keel at dissection. A total of 93 keels were correctly scored as having no fracture (True -). Accuracy of deviation detection by palpation was 91%. A total of 202 keels were assessed.

| Severity | True + | False - | | False + | |
|----------|----------|-----------|------------|-----------|------------|
| | | Palpation | Dissection | Palpation | Dissection |
| None | N/A | 17 | 0 | 0 | 2 |
| Minor | 31 (34%) | 0 | 16 (94%) | 2 (100%) | 0 |
| Severe | 59 (66%) | 0 | 1 (6%) | 0 | 0 |
| Total | 90 (45%) | 17 (8%) | | 2 (1%) | |

87%. Of the 202 keels assessed, 96 were identified as true positives and 73 as true negatives for fractures. The overall accuracy for detecting deviations by palpation was 91%, with sensitivity of 84%, and specificity 97%. Of the 202 keels assessed, 90 were identified as true positives and 93 as true negatives for deviations. The overall true prevalence of fractures and deviations determined by dissection were 53% and 56%, respectively.

In regards to fractures, 82% of false negatives were found on keels with fractures located approximately 2 cm from the caudal tip (Table 2). These fractures were likely missed by palpation since, in the majority of cases, there was no callus, bony ectostosis or suture that could be felt on the ventral surface of the *Carina sterni*. However, this type of fracture was visually detected on the dorsal surface of the keel and thus represents a category of damage which will likely be undetected by palpation independent of the level of training. Further histological analysis of this type of fracture in future studies may help provide insight into the fracture severity and healing stage of this type of injury, in particular whether this damage warrants concern. The remaining 18% of false negative fractures were on keels that had both a fracture near the tip as well as in the mid-section of the keel.

When considering the relationship between deviations and false negative fractures (i.e., fractures were present though not recorded), 60% of false negatives were from non-deviated keels. The majority of these missed fractures were from the type of keel described

above where the fracture can only be detected on the dorsal surface, i.e., not by palpation. It is possible that perfectly straight, non-deviated keels (with fractures) are occasionally assumed to be normal, non-fractured, keels when mixed in with a group of hens with highly damaged or deviated keels, as the assessor either develops a bias or does not palpate them as thoroughly. Extra caution needs to be taken to use a consistent and thorough palpation technique on every keel that is scored, particularly when prevalence of fractures and/or deviations in the population are high. The second most common source of undiagnosed fractures (23%) were found on keels with a relatively severe deviation where the *Carina sterni* manifested a folding over, i.e., the keel was no longer straight in the sagittal plane. These fractures were found on the lateral surface of the keel where the fold deviation appeared to have caused a hairline or stress fracture. Unless the fracture is severe, detecting this type of fracture on the lateral surface of the keel is very difficult by palpation alone.

Of the false positives that were detected for fractures, 64% were found in keels with a small deformation on the keel, within 0.2 cm from the caudal tip. In these cases the very caudal tip of the keel appears to have curved forward in the frontal plane sometime during calcification of the keel. This deformation, detected as a small raised structure during palpation, was misinterpreted as a callus, but no visible signs of fracture were identified during dissection.

Although scoring accuracy for deviations was high (Table 3), of the false negatives that were detected

during dissection, 94% were found on keels with only a minor deviation that may have been difficult to detect during palpation due to muscle or fat deposition. False positives were rare for deviations (1%).

One potential limitation of this case study is that all palpation scoring and dissection scoring were completed by the same observer. The errors made by this single observer may not be the same as other observers using palpation, making external validity a potential concern. Prior to scoring the keels used in this case study, the observer completed training in keel palpation scoring as an assessor in Petrik *et al.* (2013) described previously. Further practice of keel palpation was completed on 250 Lohmann LSL Lite hens ranging in age from 16 to 50 weeks with a variety of KBD, although this was done without subsequent dissection. All palpation training and practice was completed prior to the scoring reported in this case study.

Recommendations

We recommend that all assessors develop a clear understanding of the form and function of the keel bone. Assessors should be given the opportunity to palpate a wide variety of keels ranging from straight intact keels, to mildly and severely deviated or fractured keels. Following familiarization with the keel bone and its various forms, observers should practice palpation scoring followed by visual assessment during dissection. Results from Petrik *et al.* (2015) suggest that practice on a minimum of 10 and more likely 50 laying hens is required to achieve an acceptable level of accuracy and reliability in a population with a prevalence of 60 to 65%. The number of hens needed for achieving this level is likely to vary with the age and genetic line of birds, type of housing system, diet, and even management style, and may not be transferred easily if these circumstances change. It should be noted that the lower the prevalence of KBD in the population of hens used for training, the greater the number of animals required. Discussion and consensus among multiple assessors regarding the keel status and the sources of error during palpation should follow the dissections. Accuracy, sensitivity, and specificity should be calculated, and a minimum of 75%, 0.80, and 0.60, respectively, be achieved if results are to be considered valid. If 3-dimensional imaging were made available, sessions in which trainees could compare palpation to the image of the structure would be highly beneficial. When multiple assessors are involved, Kappa values and prevalence of damage should also be reported. After acceptable levels have been achieved, palpation scoring in the field can begin; however, periodic recalibration is encouraged for all examiners, regardless of experience level, in order to reduce drift over time from the standard palpation protocol. Records of these measures should be maintained and provided as assurance that KBD is being assessed accurately in deliverables such as publications.

CONCLUSIONS

The current work seeks to improve the quality of KBD detection by offering what we believe to be a simplified but superior method of categorizing and reporting keel conditions that incorporates both deviations and fractures as separately recorded characteristics. It is also necessary that the assessor improve their own abilities and thus we advocate learning from and adopting techniques from related fields which regularly use palpation in pursuit of this objective. Lastly, methods in addition to palpation are reviewed in order to provide researchers alternatives to palpation that can improve the reliability of their assessments but also offer information that cannot be achieved with palpation and/or dissection alone.

ACKNOWLEDGMENTS

This manuscript is a product of discussion and collaboration between participants of the 2014 International Keel Bone Workshop held April 2014 in Zollikofen, Switzerland. Organized by the University of Bern, funding was gratefully received from: The Swiss National Science Foundation, The Federal Food Safety and Veterinary Office of Switzerland, the Biological and Biotechnology Sciences Research Council of the United Kingdom, and the Universities Federation for Animal Welfare. Workshop participants, in alphabetical order, were: Michael Appleby (World Animal Protection), Teresa Casey-Trott (University of Guelph, Canada), Rachel Dennis (University of Maryland, USA), Ernst Fröhlich (The Federal Food Safety and Veterinary Office, Switzerland), Sabine Gebhardt (University of Bern, Switzerland), Alexandra Harlander (University of Guelph, Canada), Jasper Heerkens (Institute for Agricultural and Fisheries Research, Belgium), Patricia Hester (Purdue University, USA), Darrin Karcher (Michigan State University, USA), Ine Kempen (Proefbedrijf Pluimveehouderij, Belgium), Tine Lentfer (Bundesanstalt für Landwirtschaft und Ernährung, Germany), Maja Makagon (Purdue University, USA), Mohammed Nasr (Zagzig University, Egypt), Prafulla Regmi (Michigan State University, USA), Nadine Ringenberg (University of Bern, Switzerland), Bas Rodenburg (Wageningen University, The Netherlands), Victoria Sandilands (Scotland's Rural Cottage, Scotland), Lars Schrader (Friedrich Loeffler Institute, Germany), Markus Schwab (University of Bern, Switzerland), Janice Siegford (Michigan State University, USA), Janja Širovnik (University of Bern, Switzerland), Ariane Stratmann (University of Bern, Switzerland), John Tarlton (University of Bristol, England), Bret Tobalske (University of Montana, USA), Simi Tomar (Central Avian Research Institute, India), Michael Toscano (University of Bern, Switzerland), Stephen Turner (British Egg Industry Council, Great Britain), Teun van de Braak (Hendrix Genetics, the Netherlands), Tina Widowski (University of Guelph, Canada),

Lindsay Wilkins (Amberly Animal Welfare Consulting, England), and Hanno Würbel (University of Bern, Switzerland).

REFERENCES

- Baillie, S., A. Crossan, S. Reid, and S. Brewster. 2003. Preliminary development and evaluation of a bovine rectal palpation simulator for training veterinary students. *Cattle Pract.* 11:101–106.
- Barnett, J. L., R. Tauson, J. A. Downing, V. Janardhana, J. W. Lowenthal, K. L. Butler, and G. M. Cronin. 2009. The effects of a perch, dust bath, and nest box, either alone or in combination as used in furnished cages, on the welfare of laying hens. *Poult. Sci.* 88:456–470.
- Bossaert, P., L. Leterme, T. Caluwaerts, S. Cools, M. Hostens, I. Kolkman, and A. de Kruif. 2009. Teaching transrectal palpation of the internal genital organs in cattle. *J. Vet. Med. Educ.* 36:451–460.
- Degenhardt, B. F., J. C. Johnson, K. T. Snider, and E. J. Snider. 2010. Maintenance and improvement of interobserver reliability of osteopathic palpatory tests over a 4-month period. *JAOA J. Am. Osteopath. Assoc.* 110:579–586.
- Degenhardt, B. F., K. T. Snider, E. J. Snider, and J. C. Johnson. 2005. Interobserver reliability of osteopathic palpatory diagnostic tests of the lumbar spine: improvements from consensus training. *JAOA J. Am. Osteopath. Assoc.* 105:465–473.
- Dohoo, I. R., W. Martin, and H. Stryhn. 2009. Screening and diagnostic tests. Pages 91–133 in *Veterinary Epidemiologic Research*. 2nd ed. VER Inc, Charlottetown, PEI.
- Donaldson, C. J., M. E. E. Ball, and N. E. O'Connell. 2012. Aerial perches and free-range laying hens: The effect of access to aerial perches and of individual bird parameters on keel bone injuries in commercial free-range laying hens. *Poult. Sci.* 91:304–315.
- Einhorn, T. A. 2005. The science of fracture healing. *J. Orthop. Trauma* 19:S4–S6.
- Ericsson, K. A., M. J. Prietula, and E. T. Cokely. 2007. The making of an expert. *Harv. Bus. Rev.* 85:114.
- Esteves, J. E., and C. Spence. 2014. Developing competence in diagnostic palpation: Perspectives from neuroscience and education. *Int. J. Osteopath. Med.* 17:52–60.
- FAWC. 2010. Opinion on Osteoporosis and Bone Fractures in Laying Hens. Farm Animal Welfare Council, London.
- FAWC. 2013. An open letter to Great Britain Governments: Keel bone fracture in laying hens.
- Fleming, R. H., H. A. McCormack, L. McTeir, and C. C. Whitehead. 2004. Incidence, pathology and prevention of keel bone deformities in the laying hen. *Br. Poult. Sci.* 45:320–330.
- Gebhardt-Henrich, S. G., and E. Fröhlich. 2013. Laying rate and foot health influenced keel bone fractures in laying hens. in *Proceedings of the 9th European Symposium on Poultry Welfare*. Uppsala, Sweden.
- Habig, C., and O. Distl. 2013. Evaluation of bone strength, keel bone status, plumage condition and egg quality of two layer lines kept in small group housing systems. *Br. Poult. Sci.* 54:413–24.
- Harlander-Matauschek, A., T. B. Rodenburg, V. Sandilands, B. W. Tobalske, and M. J. Toscano. 2015. Causes of keel bone damage and their solutions in laying hens. *Worlds Poult. Sci. J.* <http://hdl.handle.net/11262/10798>.
- Heerkens, J., E. Delezie, I. Kempen, J. Zoons, T. B. Rodenburg, and F. Tuytens. 2013. Do keel bone deformations affect egg-production in end-of-lay housing hens housed in aviaries? Page 127 in 9th European Poultry Conference. R. Tauson, H.J. Blokhuis, L. Berg, and A. Elson eds. Uppsala, Sweden.
- Hester, P. Y., and S. A. Enneking. 2014. Relationship between keel mineralization and deformities of seventy-one week old White Leghorn hens. Page 51 in *Proceedings of 48th Congress of the International Society for Applied Ethology*. I. Estevez, X. Manteca, R. H. Marin, and X. Averó, eds. Wageningen Academic Publishers, Vitoria-Gasteiz, Spain.
- Hróbjartsson, A., A. S. S. Thomsen, F. Emanuelsson, B. Tendal, J. Hilden, I. Boutron, P. Ravaut, and S. Brorson. 2013. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and non-blinded assessors. *Can. Med. Assoc. J.* : cmaj-120744.
- Kappeli, S., S. G. Gebhardt-Henrich, E. Fröhlich, A. Pfulg, and M. H. Stoffel. 2011. Prevalence of keel bone deformities in Swiss laying hens. *Br. Poult. Sci.* 52:531–536.
- Knowles, T. G., and L. J. Wilkins. 1998. The problem of broken bones during the handling of laying hens—a review. *Poult. Sci.* 77:1798–1802.
- Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics.* 159–174.
- Lay, D. C., R. M. Fulton, P. Y. Hester, D. M. Karcher, J. B. Kjaer, J. A. Mench, B. A. Mullens, R. C. Newberry, C. J. Nicol, N. P. O'Sullivan, and R. E. Porter. 2011. Hen welfare in different housing systems. *Poult. Sci.* 90:278–294.
- McCormick, J. P., S. J. Morgan, and W. R. Smith. 2003. Clinical effectiveness of the physical examination in diagnosis of posterior pelvic ring injuries. *J. Orthop. Trauma* 17:257–261.
- McKenna, S. L. B., and I. R. Dohoo. 2006. Using and interpreting diagnostic tests. *Vet. Clin. North Am. Food Anim. Pract.* 22: 195–205.
- Myers, H. L., E. Thomas, E. M. Hay, and K. S. Dziedzic. 2011. Hand assessment in older adults with musculoskeletal hand problems: a reliability study. *BMC Musculoskelet. Disord.* 12:3.
- Nasr, M. A. F., J. Murrell, L. J. Wilkins, and C. J. Nicol. 2012a. The effect of keel fractures on egg production parameters, mobility and behaviour in individual laying hens. *Anim. Welf.* 21: 127–135f.
- Nasr, M. A. F., C. J. Nicol, and J. C. Murrell. 2012b. Do Laying Hens with Keel Bone Fractures Experience Pain? *PLoS One* 7:e42420.
- Nasr, M. A. F., C. J. Nicol, L. Wilkins, and J. C. Murrell. 2015. The effects of two non-steroidal anti-inflammatory drugs on the mobility of laying hens with keel bone fractures. *Vet. Anaesth. Analg.* 42:197–204.
- Nicholson, C. J., and N. E. O'Connell. 2010. Influence of access to aerial perches on keel bone injuries in laying hens on commercial free range farms. Page 250 in *Proceedings of the 44th Congress of the International Society of Applied Ethology: Coping in Large Groups*.
- Norman, S. T., and G. Dall'Alba. 2013. Computer Assisted Learning for Improving Cattle Palpation Skills of Veterinary Students. *Open J. Vet. Med.* 3:319.
- Petrik, M. T., M. T. Guerin, and T. M. Widowski. 2013. Keel fracture assessment of laying hens by palpation: inter-observer reliability and accuracy. *Vet. Rec.* 173:500.
- Petrik, M. T., M. T. Guerin, and T. M. Widowski. 2015. On-farm comparison of keel fracture prevalence and other welfare indicators in conventional cage and floor-housed laying hens in Ontario, Canada. *Poult. Sci.* 94:579–585.
- Pickel, T., L. Schrader, and B. Scholz. 2011. Pressure load on keel bone and foot pads in perching laying hens in relation to perch design. *Poult. Sci.* 90:715–24.
- Pugh, C. M. 2013. Application of National Testing Standards to Simulation-Based Assessments of Clinical Palpation Skills. *Mil. Med.* 178:55–63.
- Regmi, P., K. E. Anderson, and D. M. Karcher. 2013. Comparison of bone quality between strains and housing systems in end-of-lay hens. *Poult. Sci.* 92:83.
- Regmi, P., and D. M. Karcher. 2013. Influence of perching on keel bone deformities in laying hens. *Poult. Sci.* 92:92.
- Richards, G. J., M. A. Nasr, S. N. Brown, E. M. G. Szamocki, J. Murrell, F. Barr, and L. J. Wilkins. 2011. Use of radiography to identify keel bone fractures in laying hens and assess healing in live birds. *Vet. Rec.* 169:279.
- Richards, G. J., L. J. Wilkins, T. G. Knowles, F. Booth, M. J. Toscano, C. J. Nicol, and S. N. Brown. 2012. Pop hole use by hens with different keel fracture status monitored throughout the laying period. *Vet. Rec.* 170:494.
- Rodenburg, T. B., F. A. M. Tuytens, K. De Reu, L. Herman, J. Zoons, and B. Sonck. 2008. Welfare assessment of laying hens in furnished cages and non-cage systems: an on-farm comparison. *Anim. Welf.* 17:363–373.
- Sandilands, V., L. Baker, S. Brocklehurst, L. Toma, and C. Moinard. 2010. Are perches responsible for keel bone deformities in laying hens? Page 249 in *Proceedings of the 44th Congress of the International Society of Applied Ethology: Coping in Large Groups*. L. Lidfors, H.J. Blokhuis, and L. Keeling, eds. Wageningen Academic Publishers, Uppsala, Sweden.

- Sandilands, V., C. Moinard, and N. H. C. Sparks. 2009. Providing laying hens with perches: fulfilling behavioural needs but causing injury? *Br. Poult. Sci.* 50:395–406.
- Scholz, B., S. Rönchen, H. Hamann, M. Hewicker-Trautwein, and O. Distl. 2008. Keel bone condition in laying hens: a histological evaluation of macroscopically assessed keel bones. *Berl. Munch. Tierarztl. Wochenschr.* 121:89–94.
- Seffinger, M. A., W. I. Najm, S. I. Mishra, A. Adams, V. M. Dickerson, L. S. Murphy, and S. Reinsch. 2004. Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. *Spine (Phila. Pa. 1976)*. 29: E413–E425.
- Van Sint Jan, S. 2007. Color atlas of skeletal landmark definitions: guidelines for reproducible manual and virtual palpations. Elsevier Health Sciences.
- Stratmann, A., E. K. F. Fröhlich, S. G. Gebhardt-Henrich, A. Harlander-Matauschek, H. Würbel, and M. J. Toscano. 2015a. Modification of aviary design reduces incidence of falls, collisions and keel bone damage in laying hens. *Appl. Anim. Behav. Sci.* 165:112–123.
- Stratmann, A., E. K. F. Fröhlich, A. Harlander-Matauschek, L. Schrader, M. J. Toscano, H. Würbel, and S. G. Gebhardt-Henrich. 2015b. Soft Perches in an Aviary System Reduce Incidence of Keel Bone Damage in Laying Hens. *PLoS One* 10: e0122568.
- Tarlton, J. F., L. J. Wilkins, M. J. Toscano, N. C. Avery, and L. Knott. 2013. Reduced bone breakage and increased bone strength in free range laying hens fed omega-3 polyunsaturated fatty acid supplemented diets. *Bone* 52:578–586.
- Tauson, R., and P. Abrahamsson. 1994. Foot and Skeletal Disorders in Laying Hens: Effects of Perch Design, Hybrid, Housing System and Stocking Density [internet]. *Acta Agric. Scand. Sect. A – Anim. Sci.* 44:110–119.
- Toscano, M. J., L. J. Wilkins, G. Millburn, K. Thorpe, and J. F. Tarlton. 2013. Development of an ex vivo protocol to model bone fracture in laying hens resulting from collisions. (PE Witten, Ed.). *PLoS One* 8:e66215.
- Toscano, M. J., L. J. Wilkins, and J. F. Tarlton. 2012. Impact of a mixed chain length omega-3 fatty acid diet on production variables in a commercial free-range laying hens. *Br. Poult. Sci.* 53:360–365.
- Ullrich, S., and T. Kuhlen. 2012. Haptic palpation for medical simulation in virtual environments. *Vis. Comput. Graph. IEEE Trans.* 18:617–625.
- Viera, A. J., and J. M. Garrett. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med* 37:360–363.
- Vits, A., D. Weitzenburger, H. Hamann, and O. Distl. 2005. Production, egg quality, bone strength, claw length, and keel bone deformities of laying hens housed in furnished cages with different group sizes. *Poult. Sci.* 84:1511–1519.
- Wiest, P. W., M. F. Hartshorne, P. D. Inskip, L. A. Crooks, B. S. Vela, R. J. Telepak, M. R. Williamson, R. Blumhardt, J. M. Bauman, and M. Tekkel. 1998. Thyroid palpation versus high-resolution thyroid ultrasonography in the detection of nodules. *J. Ultrasound Med.* 17:487–496.
- Wilkins, L. J., S. N. Brown, P. H. Zimmerman, C. Leeb, and C. J. Nicol. 2004. Investigation of palpation as a method for determining the prevalence of keel and furculum damage in laying hens. *Vet. Rec.* 155:547–549.
- Wilkins, L. J., J. L. McKinstry, N. C. Avery, T. G. Knowles, S. N. Brown, J. Tarlton, and C. J. Nicol. 2011. Influence of housing system and design on bone strength and keel bone fractures in laying hens. *Vet. Rec.* 169:414.
- Wood, L., G. Peat, R. Wilkie, E. Hay, E. Thomas, and J. Sim. 2006. A study of the noninstrumented physical examination of the knee found high observer variability. *J. Clin. Epidemiol.* 59:512–520.